Bulgarian Language Journal 63 (2016), 4, 7-36
ISSN 0005-4283

Güneş Erkan
University of Michigan
Dragomir R. Radev
Yale University

LEXRANK: GRAPH-BASED LEXICAL CENTRALITY AS SALIENCE IN TEXT SUMMARIZATION

(Abstract)

We introduce a stochastic graph-based method for computing relative importance of textual units for Natural Language Processing. We test the technique on the problem of Text Summarization (TS). Extractive TS relies on the concept of sentence salience to identify the most important sentences in a document or set of documents. Salience is typically defined in terms of the presence of particular important words or in terms of similarity to a centroid pseudo-sentence.

Keywords: automatic text summarisation, lexical centrality, graph-based methods

✉ Güneş Erkan
gerkan@umich.edu

✉ Dragomir R. Radev
dragomir.radev@yale.edu